ABSTRACT

Real Time framework for Parallel streaming data Processing framework (RT-ParaPro) is a framework used to develop programs that simultaneously process streaming data transmitted over a real-time network and send data over a network and archive them in real time.

• LHML, which determines whether plasma is L-mode or H-mode in real time by using machine learning, and RFM Archiving system, which stores various RFM channel data to Mdsplus, has been developed and operated in KSTAR.

• The period of the thread in the RT-ParaPro has a jitter of about 8 usec not only in the low control cycle rate (1kHz) but also in the control cycle rate of 100kHz.

BACKGROUND

• Most real-time frameworks (such as ITER RTF, MARTe) were developed with real-time control in mind, not real-time parallel streaming data processing for real time network.

• Most stream processing framework (Such as Spark Streaming, Flink, Storm) support millisecond(msec) order latency.

• RT-ParaPro is specialized in the parallel data processing, archiving and transmission of data over a real-time network and support microsecond(usec) order latency.

CHALLENGES / METHODS / IMPLEMENTATION

Challenges

In KSTAR, 50 real time system connected to real time network(such as SDN, RFM) to process data in real time.

- These nodes process the data transmitted in multiple data channels in parallel.

- In KSTAR, the system with the highest number of channels must simultaneously process streaming data from 158 different channels in parallel in 2 kHz period.

Methods/Implementation

This framework consists of pairs of thread and buffer which implements parallel producer/consumer design pattern(C++).

- Each thread is able to set the attributes needed to have real-time properties (affinity, period, etc) by using RT-ParaPro.
- Control thread controls each thread in accordance with user command or shot sequence.

OUTCOME

Applications

By using RT-ParaPro, 2 applications are developed in KSTAR

• LHML : this application determines whether plasma is L-mode or H-mode in real time by using machine learning(LSTM). (RT threads running at 1kHz, LSTM takes 56 μsec)

• RFM Archiving system : this application stores 158 RFM channel data to Mdsplus, has been developed and operated in KSTAR. (RT threads running at 1kHz)

Performance Test

To evaluate the real time performance of this framework, we tested the consistency of thread period by varying the period of the thread (1kHz,2kHz,5kHz,10kHz,100kHz).

- The test shows that the thread control period is consistent.
- The period of the thread has a jitter of about 8 usec not in the low control cycle rate (1kHz) but also in the control cycle rate of 100kHz.

Table of consistency of thread period

<table>
<thead>
<tr>
<th>Thread Period</th>
<th>Mean(nsec)</th>
<th>Std(nsec)</th>
<th>Max - Min(nsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1kHz</td>
<td>999999.996</td>
<td>176.9456</td>
<td>8369.0000</td>
</tr>
<tr>
<td>2kHz</td>
<td>999999.994</td>
<td>171.8813</td>
<td>6654.0000</td>
</tr>
<tr>
<td>5kHz</td>
<td>199999.994</td>
<td>169.1178</td>
<td>8528.0000</td>
</tr>
<tr>
<td>10kHz</td>
<td>999999.997</td>
<td>171.5877</td>
<td>7620.0000</td>
</tr>
<tr>
<td>100kHz</td>
<td>999999.995</td>
<td>169.9920</td>
<td>7325.0000</td>
</tr>
</tbody>
</table>

(Left) Consistency of RT thread period at various threads frequency.

(Right) Small latency spike(around 3000 nsec) was presented every 100 counts.

CONCLUSION

• In this paper, we present a Real Time framework for Parallel streaming data Processing framework (RT-ParaPro).

• This framework is specialized in the parallel data processing, archiving and transmission of data over a real-time network. By using this framework, KSTAR developed two real-time data processing applications which need to have microsecond(usec) order latency.

• Periodical latency spike problems will be solved.

• GPGPU processing module and real-time image data processing module are under development.

ACKNOWLEDGEMENTS / REFERENCES